Structure, Vol. 9, 409–417, May 1, 2001, ©2001 Elsevier Science Ltd. All rights reserved. PII S0969-2126(01)00601-3

Structure of Cry2Aa Suggests an Unexpected Receptor Binding Epitope

R.J. Morse,[1] T. Yamamoto,[2,4] and R.M. Stroud[1,3]
[1] Department of Biochemistry and Biophysics
University of California, San Francisco
San Francisco, California 94143
[2] Sandoz Agro
975 California Avenue
Palo Alto, California 94304

Summary

Background: Genetically modified (GM) crops that express insecticidal protein toxins are an integral part of modern agriculture. Proteins produced by *Bacillus thuringiensis* (Bt) during sporulation mediate the pathogenicity of Bt toward a spectrum of insect larvae whose breadth depends upon the Bt strain. These transmembrane channel-forming toxins are stored in Bt as crystalline inclusions called Cry proteins. These proteins are the active agents used in the majority of biorational pesticides and insect-resistant transgenic crops. Though Bt toxins are promising as a crop protection alternative and are ecologically friendlier than synthetic organic pesticides, resistance to Bt toxins by insects is recognized as a potential limitation to their application.

Results: We have determined the 2.2 Å crystal structure of the Cry2Aa protoxin by multiple isomorphous replacement. This is the first crystal structure of a Cry toxin specific to Diptera (mosquitoes and flies) and the first structure of a Cry toxin with high activity against larvae from two insect orders, Lepidoptera (moths and butterflies) and Diptera. Cry2Aa also provides the first structure of the proregion of a Cry toxin that is cleaved to generate the membrane-active toxin in the larval gut.

Conclusions: The crystal structure of Cry2Aa reported here, together with chimeric-scanning and domain-swapping mutagenesis, defines the putative receptor binding epitope on the toxin and so may allow for alteration of specificity to combat resistance or to minimize collateral effects on nontarget species. The putative receptor binding epitope of Cry2Aa identified in this study differs from that inferred from previous structural studies of other Cry toxins.

Introduction

The almost 20 million hectares of GM crop fields in North America consist of crops engineered for herbicide or insect resistance. The genes that confer the latter trait come from *Bacillus thuringiensis* (Bt), a family of Gram-positive sporulating soil bacteria that produce parasporal crystals with insecticidal activity. The insecticidal activity of particular Bt isolates is directed against narrow spectra of insect larval species, usually within a single order. Bacterial toxins known as insecticidal crystal proteins (ICPs) or crystalline (Cry) proteins that are sequestered as protoxins in crystalline inclusions after sporulation mediate this species-specific pathogenicity [1]. The Cry protoxins are ingested, solubilized in the larval gut [2, 3], and activated by the removal of an amino-terminal segment and a C-terminal segment, the size of which depends on the gene or its protoxin [2, 4]. The active toxins associate with insect-specific receptors of gut epithelial cells of the target insect [5] and subsequently insert into the cell membrane [6, 7], leading to the formation of ion channels [8, 9, 10]. This results in disruption of the electrochemical balance across the basal membrane, gut paralysis, and larval death [11, 12, 13, 14]. The host cadaver serves as growth medium for vegetative cells arising from germination of the Bt endospores.

Species selectivity of Cry proteins is encoded in the binding site for the target receptor [5]. Classification of the Cry proteins is based on amino acid sequence identity [15] and is roughly correlated with the taxonomic order of susceptible insect species, spanning species of agricultural (Cry1 Lepidoptera, Cry2 Lepidoptera, and Cry3 Coleoptera) and public health (Cry2 and Cry4 Diptera) significance. The structure may help guide mutagenesis followed by screening that is directed toward the fine tuning of species selectivity in order to design insecticides that do not kill nontarget organisms such as monarch larvae [16]. It also may assist in the elucidation of the structural basis of resistance to Bt toxins and the subsequent generation of novel insecticidal toxins for use on Bt-resistant insects [17, 18].

Structure-based protein engineering of Cry toxins may direct the search for variants with broader susceptible species spectra, optimal potency, and stability properties. Cry2Aa is among an unusual subset of Cry proteins possessing broad insect species specificity by exhibiting high specific activity against two insect orders, Lepidoptera and Diptera [19, 20]. It is lethal to more lepidopteran species than the Cry1 toxins deployed against agriculturally important Lepidoptera [21] and exhibits a low level of crossresistance in Cry1A-resistant insects [22]. Also, the mode of action of Cry2Aa may be distinct from that of other Cry toxins [23]. Thus, it could serve as a platform for the design of Cry toxins with broader susceptible species spectra and minimal Cry1A-derived crossresistance in the field.

Chimeric-scanning mutagenesis experiments have identified disjoint blocks (D and L, see Results and Discussion) of the Cry2Aa sequence that separately confer specificity against dipteran (D) and lepidopteran (L) species [24, 25]. These experiments also demonstrate that maximal activity against lepidopteran species requires not only L block residues but also some of the specificity determinants of the D residue block. Further, Cry2Ab, an 87% sequence identical homolog of Cry2Aa, has

[3] Correspondence: stroud@msg.ucsf.edu
[4] Present address: Maxygen, Redwood City, California, 94063.

Key words: Bacillus thuringiensis; delta-endotoxin; Cry2Aa; binding epitope; crystal structure; X-ray

Table 1. Data Collection, MIR Phasing, and Structure Refinement Statistics for Cry2Aa

	Native	$U(NO_3)$	PtCNS	PtI_6	$NbCl_2$	Ru^1	Hg^2
Data Collection (1.08 Å)							
Resolution (Å)	2.2	2.6	2.5	2.6	2.6	2.5	2.5
Unit cell dimensions (Å)	a = b = 85.6						
	c = 163.9						
Space group	$P4_32_12$						
Number of observed reflections (σ_F – 2.5)	245,580	69,703	139,057	70,618	73,949	113,242	126,930
Number of unique reflections	31,591	17,370	20,476	17,999	17,455	19,475	20,198
Completeness (%)	99.3	89.1	99.0	92.	89.3	94.7	97.0
R_{merge} (%)	5.7	5.1	5.7	4.7	4.3	5.4	6.3
Phasing/MIR							
Resolution		2.6	2.6	2.6	4.5	2.6	5.25
Number of sites refined		5	6	7	6	5	3
Number of reflections (σ_F 3)		16,177	17,808	16,516	3,136	17,074	2,419
R_{iso} (%)		16	24	32	10	10	16
R_{cullis}		.62	.62	.59	.60	.67	.62
R_{kraut}		.13	.15	.20	.06	.08	.09
Phasing power		1.1	1.9	1.8	1.4	0.8	1.2
$<FOM>_{centric}$[3]		.36	.39	.41	.41	.30	.42
$<FOM>_{overall(n_{phased})}$[4]	.65 (18,677)						
Refinement							
Resolution (Å)	28.0–2.2						
Number of reflections (completeness %)	31,509 (93)						
R_{cryst} [σ_F = 0] (2.3–2.2 Å)	.18 (.21)						
R_{free} [5% test] (2.3–2.2 Å)	.24 (.23)						
Number of non hydrogen atoms	5,001						
Number of water molecules	514						
Rms bond distances (Å)	.005						
Rms bond angles (°)	1.2						

[1] $[Ru(NH_3)_6]Cl_3$.
[2] Para chloromercuri phenol (PCMP).
[3] Individual data set results.
[4] Final number of phased reflections.

negligible activity against dipteran species and 3- to 8-fold less activity against certain lepidopteran species [25, 26]. Hence, Cry2Aa structure and mutagenesis data provide the basis for future protein engineering of Cry toxins with modified specificity and selectivity profiles.

To understand the structural determinants of Cry toxin specificity, we determined the crystal structure of the protoxin of Cry2Aa from *Bacillus thuringiensis* subsp. *kurstaki*. The complete structure was determined by multiple isomorphous replacement and refined to 2.2 Å resolution. We have identified a candidate toxin receptor binding surface that is consistent with available chimeric-scanning mutagenesis data.

Results and Discussion

The structure of Cry2Aa from *Bacillus thuringiensis* subsp. *kurstaki* was determined by multiple isomorphous replacement using six heavy atom derivatives and was refined to 2.2 Å resolution with R_{cryst} = 18% (Table 1). The structure of the 633-amino acid protoxin contains the N-terminal 49-amino acid peptide that is cleaved upon activation and the three domains of what will become the mature toxin [27]. The structures of the three domains are surprisingly similar in overall topology (Figure 1a) to those of the activated toxins Cry3Aa [28] and Cry1Aa [29], suggesting that removal of the activation peptide serves to expose regions of the toxin rather than alter its conformation. This structural homology is also surprising since these toxins have little sequence identity to Cry2Aa (20% and 17%, respectively). In the mature toxin, the N-terminal domain (residues 1–272) is a pore-forming seven-helical bundle (Figure 1d) [1]. The second domain (residues 273–473) is a receptor binding β prism, a three-fold symmetric arrangement of β sheets, each with a Greek key fold (Figure 1e). The third domain (residues 474–633) is implicated in determining both larval receptor binding [30, 31] and pore function [32] and is a lectin-like C-terminal β sandwich (Figure 1f).

Available chimeric-scanning mutagenesis data [24, 25] define a candidate toxin-receptor binding surface on Cry2Aa that is comprised of a distribution of hydrophobic residues (Ile474–Ala477 from β12a, Val365–Leu369 from the β5-β6 loop, and Leu402–Leu404 from the β7-β8 loop) across the solvent-exposed surface of the β prism and β sandwich domains (Figure 1b). Proteolytic activation of the toxin involves the removal of the 49 N-terminal amino acids and exposes residues comprising this putative toxin-receptor binding surface. Removal of the 49 amino terminal residues, comprised of α0, α0a, and an N-terminal coil, would not affect the structure of the seven-helical membrane insertion domain, as seen by comparing the structures of the activated toxin Cry1Aa and that of the protoxin Cry2Aa.

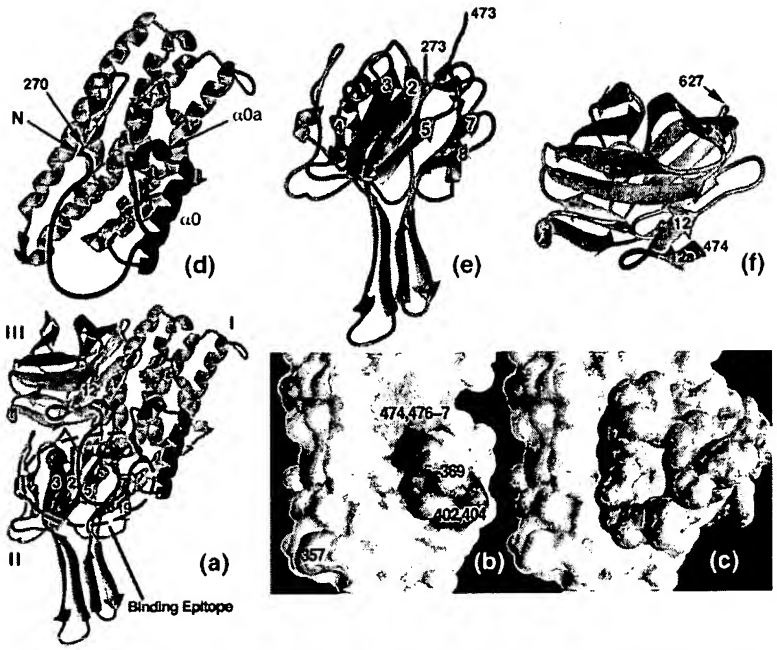

Figure 1. Topology and Solvent Accessible Surface of Cry2Aa

(a) Ribbon diagram, rendered by Midas Plus [48], of Cry2Aa. Domain I is shown in magenta, domain II is shown in blue, and domain III is shown in cyan. The N terminus is shown in red, while functionally important loops delimiting the putative toxin-receptor binding epitope are shown in green. A Cry2Aa insertion, relative to Cry3Aa and Cry1Aa, before β12 at the N terminus of domain III is shown in magenta. Numbered β strands referred to in the text are labeled.

(b) The solvent accessible surface, as calculated by GRASP [49], of domains II and III of Cry2Aa. The orientation is identical to that shown in Figure 1a. The projection of residue hydrophobicity onto this surface is shown in color. Portions of the hydrophobic surface contributed by residues 474, 476, and 477 are shown in cyan, those contributed by residues 365–369 are shown in blue, those contributed by residues 402 and 404 are shown in magenta, and the remainder of the surface contributed by hydrophobic residues is shown in yellow. The remaining surface that is identified as nonhydrophobic is colored white. Residue hydrophobicity is as defined by GRASP [49]. The prominent hydrophobic patch is the center of the putative toxin-receptor binding epitope. For orientation, the portion of the surface contributed by residue 357 of the β4-β5 loop is shown in red.

(c) The solvent accessible surface (as calculated by GRASP) of domains II and III of Cry2Aa. The orientation is identical to that shown in Figure 1a. The projection of residue hydrophobicity onto this surface is shown in yellow, while the N terminus is shown in red; the N terminus sterically hinders access to the putative toxin-receptor binding epitope. Portions of the surface that are identified as nonhydrophobic are colored white.

(d–f) The three domains of Cry2Aa shown in the same orientation as in Figure 1a. Labels with amino acid numbers identify the visible N and C termini of each domain in the figures.

This is also expected since constructs consisting of the N-terminal-helical domain of the complete Cry3Ba1 protoxin (prior to cleavage) are capable of nonreceptor-mediated partitioning into lipid bilayers [33], as is the activated toxin.

The structure of Cry2Aa suggests that the N-terminal residues should sterically hinder access to the putative binding epitope β5-β6 and β7-β8 loops (Figure 1a, shown in green) and the exposed parts of domain III closest to domain II. Projection of hydrophobicity onto the solvent accessible surface of domains II and III reveals an 800 Å^2 hydrophobic patch (Figure 1b) proximal to these loops. However, while the structure suggests that the 49 N-terminal residues (α0, α0a, and the N-terminal coil) should sterically hinder access to the putative binding epitope, the biological rationale for this function is unclear. It is unlikely that Bt possesses a receptor with affinity for the activated toxin. Hence, it does not seem likely that the N terminus serves to prevent premature activation of the toxin within Bt. One simple explanation is that occlusion of the hydrophobic patch of the putative binding epitope prevents nonspecific aggregation of the toxin with itself or other host proteins. Another explanation is that the N-terminal amino acids play a role in the formation of the environmentally stable crystalline inclusions.

The specificity-distinguishing residues are also indicated by comparison of the Cry2Aa structure with the structure of the highly homologous (87% sequence identity) Cry2Ab that is inactive against some Cry2Aa target species (Figure 2a). Chimeric-scanning mutagenesis [24, 25] defines a continuous 106 amino acid block, 307–412, of specificity-distinguishing residues. (Specifically, [25] demonstrated that substitution of residues 278–340 resulted in loss of dipteran-specific activity in Cry2Aa, while [24] demonstrated that substitution of residues 307–382 conferred dipteran-specific activity to Cry2Ab. Thus, in our discussion, we adopt residue 307 as the N-terminal boundary of the specificity-conferring sequence in Cry2Aa.) Within these 106 amino acids, there are 23 residues that differ between Cry2Aa and Cry2Ab (sequence alignment presented in Figure 5). Most of the Cry2Aa-Cry2Ab amino acid differences lie within or about the domain II/III 800 Å^2 hydrophobic patch (Figure 1b) and surrounding residues from the β5-β6, β7-β8, and β4-β5 loops (Figure 1a). The picture of the putative toxin-receptor binding surface that emerges is that of an 800 Å^2 hydrophobic region surrounded by three loops, those joining β4-β5, β5-β6, and β7-β8, which are also a part of the putative binding site. The three loops contain hydrophilic side chains that may be involved in specific hydrogen bonding with the receptor and so signal a portion of the site that could be mutated both to probe these interactions and to alter specificity.

The proximity of this surface to solvent-exposed loops of the lectin-like domain III is consistent with the finding that domain III plays a role in the fine tuning susceptibility of different species. This has been demonstrated by replacement of domain III [30, 31] to make chimeric toxins with altered specificity characteristics. The N-ter-

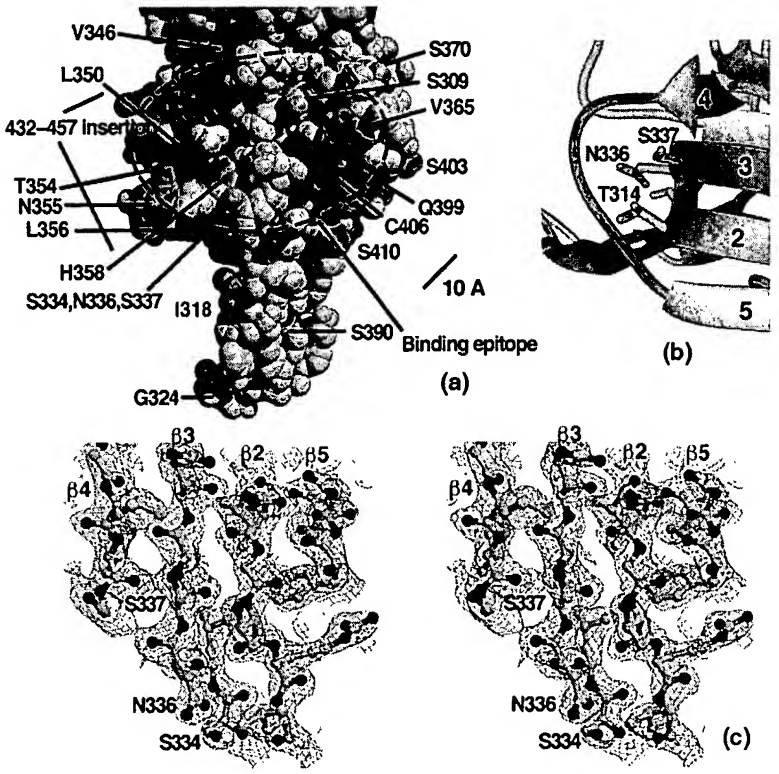

Figure 2. Space-Filling Representation of Cry2Aa Specificity-Conferring Residues, Detail of Buried D Block Residues, and Electron Density Covering Buried D Block Residues

(a) Space-filling model of Cry2Aa domains II and III with the N terminus and membrane-inserting domain I removed. The orientation reflects a −20° rotation relative to that shown in Figure 1a. The results of chimeric-scanning [24, 25] mutagenesis experiments are projected onto the van der Waals surface of Cry2Aa. The residues colored green and cyan are single amino acid differences between Cry2Aa and Cry2Ab in block L (residues 341–412). The residues colored yellow and orange are single amino acid differences between Cry2Aa and Cry2Ab in block D (residues 307–340). The bar represents an approximate 10 Å scale.

(b) Packing of D block residues behind the β4-β5 loop. The β4-β5 loop contains L block specificity determinants with which the buried D block residues interact.

(c) Electron density for the putative receptor binding site covering residues of the β sheet behind the β4-β5 loop.

minal strand β12a of domain III is not present in the three-dimensional structures of Cry1Aa or Cry3Aa. The turn between this strand and the rest of domain III is functionally replaced almost exactly by a loop that connects β3 and β4 of domain II in the homologous Cry1Aa and Cry3Aa structures (Figures 1a and 3, shown in magenta). This functionally conserved β12a motif occupies the same region of the structure as the β3-β4 turn in Cry1Aa and Cry3Aa, so it implies conservation of a functional role in protecting the hydrophobic portion of the putative receptor binding surface implicated by the homolog substitutions.

Chimeric-scanning mutagenesis identifies fairly large

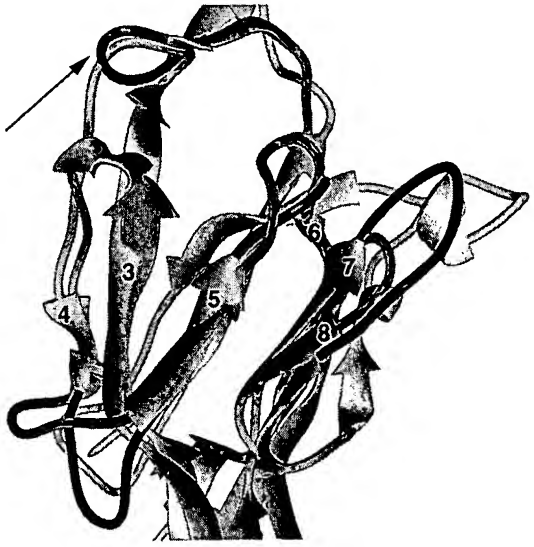

Figure 3. Detail of Ribbon Diagram Overlap of Cry2Aa and Cry1Aa

The Cry1Aa domains have been independently fit to those of Cry2Aa. The functionally important loops delimiting the putative toxin-receptor binding epitope are shown in green (Cry2Aa) and blue (Cry1Aa). The Cry2Aa insertion, relative to Cry3Aa and Cry1Aa, before β12 at the N terminus of domain III is shown in magenta, while the corresponding loop from Cry1Aa is shown in cyan (see arrow).

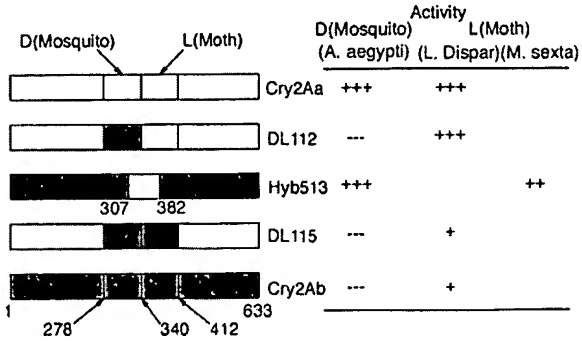

Figure 4. Schematic Representation of Chimeric-Scanning Mutagenesis Data

The first and last bands represent the Cry2Aa and Cry2Ab sequences, respectively. The middle bands represent chimeric combinations in which gray regions correspond to Cry2Ab sequence and white regions correspond to Cry2Aa sequence. For all bands, except that corresponding to Hyb513, the three central vertical bars represent amino acids 278, 340, and 412. For Hyb513, the two central vertical bars represent amino acids 307 and 382. The activity designations represent an approximate log scale. For example, the (+++) activity designation for chimera DL112 corresponds to an ID_{50} of 126 (85.7–187) ng, while the (+) designation for chimera DL115 corresponds to an ID_{50} of 3,200 (1,340–51,900) ng; the confidence intervals correspond to 2σ.

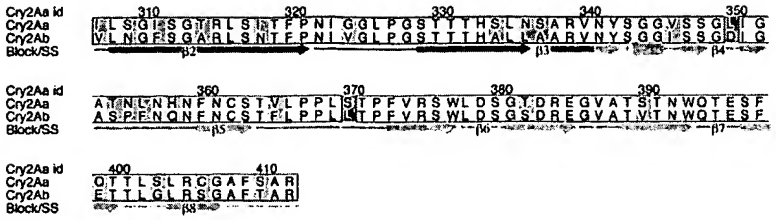

Figure 5. Detail Sequence Alignment of Cry2Aa and Cry2Ab

Sequence alignment of the D and L block regions of Cry2Aa and Cry2Ab generated using ALSCRIPT [51]. In the alignment, identical amino acids are unmarked; similar residues (as defined by ALSCRIPT) are colored yellow, while dissimilar residues are marked green. The secondary structure associated with sequence is presented in the lowermost row. The block of secondary structure associated with D block residues is colored magenta, while that associated with L block residues is colored cyan.

regions of the protein sequence that confer differential specificity to Diptera and Lepidoptera [25] (Figure 4). In Figure 4, the first band represents the sequence of Cry2Aa with its high level of activity (+++) against both Lepidoptera and Diptera. The last band represents the Cry2Ab sequence that exhibits negligible activity (---) against Diptera and up to one order of magnitude lower activity against Lepidoptera when compared with Cry2Aa. The second band (DL112) represents a Cry2Aa chimera that contains the Cry2Ab sequence for the block D residues 307–340 (dipteran-specific). This chimera has negligible activity against Diptera and is suggested to have reduced activity (at the 1σ level) against Lepidoptera, indicating that block D correlates with dipteran specificity. The activity profile of a reverse chimera (the third band) [24], in which Cry2Ab contains the block D sequence from Cry2Aa, shows a more significant reduction than DL112 against Lepidoptera (of a different species) but is only reduced 20-fold toward Diptera versus Cry2Aa. Thus antidipteran activity tracks with the D block of Cry2Aa.

The fourth band (DL115) represents a Cry2Aa chimera that contains the Cry2Ab sequence for the dipteran-disfavoring D block and for a neighboring region of sequence, the lepidopteran-disfavoring L block (residues 341–412). The activity profile of this construct against both Diptera and Lepidoptera most closely parallels that of Cry2Ab, which is consistent with blocks D and L encoding essentially all of the differential specificity determinants. In summary, the differential specificity for Diptera in Cry2Aa depends on block D, while that for Lepidoptera depends on block L. Maximal activity

Table 2. Solvent Accessible Surface Areas, Contacts within 3.4 Å, and Hydrogen Bonds for the Specificity-Conferring Residues in Cry2Aa

Residue	Exposed Surface (Å²)	Exposed Surface Beyond C_β (Å²)	Contacts
			Dipteran Specificity-Conferring
Ile307	6	4	Ser309,Ser343,Gly481,(Met483),(Tyr342)
Ser309	26	26	Asn341,Ile307,Thr364,(Ser363)
Ile311	1	0	Cys362,(Arg339),(Asn361)
Thr314	7	7	Ser337,Asn357,Asn336,His358,(Asn359)
Ile318	91	89	Thr332,(Thr331)
Gly324	78	0	
Ser334	5	5	Leu316,Asn336,(Phe409),(Gln399),(Arg315)
Asn336	6	6	Thr314,Ser334,Ala460,Ala353,(Gly313),(Ile351)
Ser337	0	0	Thr314,Ala353,His358
			Lepidopteran Specificity-Conferring
Val346	39	34	Tyr342,(Asn303),(Gly344)
Leu350	27	26	Asn449,Ile450
Thr354	50	26	Glu451
Asn355	109	76	(Pro457)
Leu356	60	43	(Ala353)
His358	43	14	Ser312,Thr314,Ser337,(Gly313)
Val365	107	75	(Asn336)
Ser370	68	39	Pro367,(His21)
Thr382	54	24	(Asn392),(Thr391)
Ser390	9	9	Ser329,Thr331,Asp383
Gln399	33	33	Val374,Arg375,Arg405,(Leu404)
Ser403	93	73	
Cys406	37	27	Ser397,Phe398,Cys362
Ser410	89	72	

All data were calculated for the activated toxin using HBPLUS [50]. Entries in the left-most column are the 23 specificity-conferring residues. Entries in the right-most column conform to hydrogen bonding geometry, except for those enclosed in parentheses that are van der Waals contacts. Bold entries in the right-most column identify specificity-conferring residues also found in the left-most column.

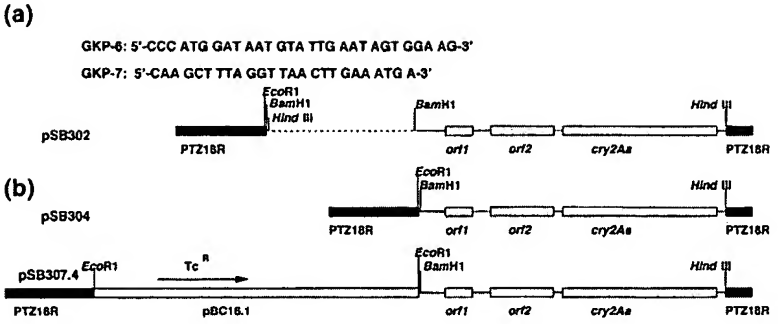

Figure 6. Restriction Maps Detailing the Construction of Plasmid pSB307

(a) Nucleotide sequences of the oligonucleotides GKP-6 and GKP-7.

(b) Restriction maps of pSB302, pSB304, and pSB307.

against Lepidoptera, as seen in Cry2Aa, still requires some contribution from block D (sequence alignment presented in Figure 5).

Figure 2a projects the Cry2Aa/Cry2Ab homolog differences onto the van der Waals surface of Cry2Aa (for clarity, only domains II and III are shown). In the D block, there are nine residues that differ between Cry2Aa and Cry2Ab. Surprisingly, most of these are buried. The notable exceptions are Ile318 and Gly324 (Asn and Val, respectively, in Cry2Ab), which are distant from the putative binding epitope, and the moderately exposed Ser309 (Asn in Cry2Ab) within the putative binding epitope (Table 2). Ile307 and Ile311 are found packed behind exposed residues on the putative binding surface. Almost half of the variant residues from block D (Thr314, Ser334, Asn336, and Ser337) are in a cluster that is packed behind the β4-β5 loop presented from within the 72-residue L block (Figures 2b and 2c).

Two of these buried variant residues, Thr314 and Ser337, make side chain-main chain hydrogen bonds with the β4-β5 loop. A third residue, Asn336, makes main chain-main chain hydrogen bonds with the β4-β5 loop, and Thr314 makes side chain-side chain hydrogen bonds with Ser334. In the less active homolog, Cry2Ab, these residues are replaced with approximately isosteric nonhydrogen bonding residues, suggesting that this pattern of substitutions abolishes affinity for the dipteran receptor (Thr314Ala, Ser334Ala, Asn336Leu, and Ser337Ala). It is conceivable that the Ile318Val and Gly324Val substitutions are part of a region of the protein that interacts only with the receptor(s) found in dipteran species and shares some components with the putative binding epitope that we identify. However, we speculate that the same exposed surface area binds to the lepidopteran and dipteran receptors. In this model, these solvent inaccessible residues behind the putative receptor binding surface may serve to alter the conformation of the β4-β5 loop, with its several hydrophilic specificity-determining residues. Similar modulation of specificity in protein-protein interactions by noncontact residues is seen in the context of immunoglobin residues that affect conformation of the complementarity-determining residues (CDR) at the binding surface [34]. Likewise, affinity maturation of a Fab/antigen complex results in the optimization of antibody/antigen binding by residues 15 Å from the interaction surface [35].

The structures of Bt toxins provide a template for design and discovery of changes that alter receptor targeting in order to either broaden selectivity for better field efficacy, prolong the life of existing agents, or avoid unwanted effects on nontarget organisms. Resistance to Bt toxins is recognized as a potential limitation in their application. Early studies concluded that recessive genes controlled the inheritance of Bt resistance. However, a recent study suggests that Bt resistance can be inherited as an incompletely dominant autosomal gene [36]. The authors note that such a mechanism of Bt resistance inheritance in the field would significantly reduce the usefulness of the high dose/refuge strategy of resistance management in which some mates are not challenged with toxin. Knowledge of any presumed modifications in the receptor that cause resistance can potentially instruct rational protein engineering of the receptor binding surface to yield toxins that might bypass resistance and still bind to the modified receptor of resistant insect species.

Potential collateral effects upon nontarget insect species [36] and effects upon nontarget predatory insects that consume target insect species [37] have been attributed to Bt GM crops. The structures provide a blueprint for focused mutagenesis followed by screening to select for each specific target species in a particular crop, so as to diminish collateral toxicity to nontarget species. By shedding light on the molecular basis of toxin-host receptor recognition, the structure provides a foundation for engineering Bt-based toxin genes that may develop broader insect species specificity, species selectivity tuned to reduce collateral impact upon nontarget species, and longer field efficacy.

Biological Implications

We have determined the three-dimensional structure of the insecticidal toxin Cry2Aa in order to understand the structural determinants of toxin specificity. Genetically modified (GM) crops that express insecticidal protein toxins are an integral part of modern agriculture. Proteins normally produced by different strains of *Bacillus thuringiensis* (Bt) during sporulation mediate a species-specific pathogenicity of Bt toward insect larvae of the target species and are the active agents in the majority of biorational pesticides and insect-resistant transgenic crops. Though promising as a crop protection alternative, problems exist with transgenic crops. Bt GM crops may pose a threat to nontarget insect species [16] or to nontarget predatory insects that consume target insect species [37]. In addition, resistance to Bt toxins is recognized as a potential limitation to their application that is ecologically friendlier than traditional organic pesticides. For instance, EPA approval of Bt GM maize was

contingent upon the establishment of viable resistance management strategies [36].

Cry2Aa is among an unusual subset of crystalline (Cry) proteins possessing broad insect species specificity by exhibiting high specific activity against larvae from two insect orders, Lepidoptera and Diptera [24, 25], of agricultural and public health significance. Also, the Cry2Aa protoxin is significantly smaller (72 kDa) than those of the Cry1 proteins (~135 kDa) in the current generation of transgenic crops. Since gene size can be a limiting factor of protein expression in plants, transgenic constructs based upon Cry1 usually express a smaller portion of the gene that contains essentially the activated toxin. Cry protoxins are presumed to be more environmentally stable than the activated toxins; hence, transgenic constructs that express the Cry2Aa protoxin could deliver higher toxin doses in the field due to greater stability [22]. Also, expression of the protoxin reduces collateral damage to nontarget insect species since it depends on specificity of the host proteases for activation [3, 37]. Chloroplast-directed overexpression of the Cry2Aa protoxin has been demonstrated and shows expression levels equivalent to 2%–3% total soluble protein in transformed leaves [22]. Such high levels of expression, 20- to 30-fold higher than current nuclear transgenics, could diminish the opportunity for developing resistance by significantly increasing toxin dose at the initial encounter with the insect.

Cry2Ab, an 87% sequence identical homolog of Cry2Aa, has negligible activity against dipteran species and 3- to 8-fold less activity against certain lepidopteran species [25, 26]. Also, there exists a unique body of chimeric-scanning mutagenesis data in the Cry2Aa/Cry2Ab system that has identified determinants of species specificity in the amino acid sequence [24, 25]. Correlating the structure with chimeric-scanning data indicates that the putative receptor binding epitope of Cry2Aa lies on the core β sheet and differs from the end of the β sheet apical loops of domain II, as suggested from structures of the other Cry toxins [28, 29]. Thus, a target surface is defined for directed mutagenesis that may focus engineering of the toxin either to develop broader insect species specificity, species selectivity tuned to reduce collateral impact upon nontarget species, or longer field efficacy. Until now, the search for new insecticidal bacterial toxins involved collection and assay of novel isolates of Bt and other bacteria known to have insecticidal activity. Recent reports describe the isolation of bacterial species that produce new classes of insecticidal toxin [38]. These structure data may permit rational engineering of insecticidal Cry toxins with desired characteristics.

Experimental Procedures

Cloning of Cry2Aa

Oligonucleotide primers flanking the coding region of *cry2Aa* were generated based on the published sequence of the gene from Bt *kurstaki* strain HD-1 [26]. Primer GKP-6 is a 29-mer that corresponds to the N-terminal 26 nucleotides of the coding region (Figure 6a). Primer GKP-7 is a 25-mer that corresponds to a fragment overlapping the HindIII site that is located ~350 nucleotides downstream from the stop codon (Figure 6a). Plasmid DNA isolated from Bt *kurstaki* HD-1 served as a template for the PCR reaction. The resulting 2100 bp fragment was purified and served as the probe used to identify the Cry2Aa operon with its accompanying open reading frames. The hexamer-primed labeling method was used to incorporate ^{32}P-dCTP into the probe.

Previously, it was indicated that the entire gene, including the coding region and the promoter, is present on a 5.0 kb HindIII fragment [26] of a plasmid isolated from Bt *kurstaki* HD-1. The 3.5–7 kb fragments obtained by HindIII digestion of plasmid DNA isolated from Bt *kurstaki* HD-1 were ligated into an *E. coli* cloning vector, pTZ18R (Pharmacia, vecbase accession #VB0071) and were used to transform *E. coli* DH5 cells by electroporation. Electroporated DH5 cells were plated onto LB-Amp^{50} plates containing X-gal and IPTG for color selection. The presence of the *cry2Aa* gene in the transformed colonies (white) was confirmed by hybridization of the PCR-generated probe. Restriction analysis was used to confirm that the clones contained inserts with the *cry2Aa* gene and also to establish the orientation with which the fragment was inserted into pTZ18R. The results of this analysis revealed that one of the clones corresponded to the orientation designated pSB302 (Figure 6b), while two clones had the opposite orientation and were designated pSB303. pSB304, obtained by deleting the 1.2 kb-BamHI fragment (dotted line in Figure 6b), was also transformed into DH5.

Total protein analysis for proteins produced by *E. coli* strain DH5 carrying pSB302, pSB303, or pSB304 was performed by SDS-PAGE. A protein band of molecular weight 62 kDa, absent in the original DH5 cells, was observed in all of the clones examined. The level of expression was the highest in those cells bearing pSB304. Most of the toxin could be found in the pelletable fraction following sonication of the cells. Samples were evaluated for biological activity by bioassay using *Manduca sexta* as the target insect. All of the clones (pSB302, pSB303, and PSB304) were active with LD_{50} values of ~500 ppm.

The pSB304 plasmid retains a unique EcoR1 site, ~200 nucleotides upstream of the *cry2Aa* promoter, into which the EcoR1-linearized *Bacillus cereus* vector pBC16.1 (GenBank accession number U32369) was cloned (Figure 6b). The resulting clone was used to transform *E. coli* DH5, and clones containing the new plasmid were designated pSB307. Confirmation of the identity of the new plasmid and determination of the orientation of the pBC16.1 insert, with respect to the *cry2Aa* gene, was made by restriction mapping. One of the plasmids, pSB307.4, was transformed into Bt *cryB* (a cry$^-$ strain) by electroporation. The plasmid content of these isolates was verified by restriction mapping.

Cry2Aa expressed well in Bt cryB cells transformed with pSB307.4, and the protein formed crystalline (rhombohedral) inclusions. The cells were harvested by centrifugation, washed with water, and lyophilized. Dried cell mass was added to the insect diet and fed to *M. sexta* larvae. The results confirmed that Bt *cryB* (pSB307.4) exhibited high insecticidal activity.

Protein Expression and Purification

The plasmid (pSB307.4) containing the Cry2Aa operon, with its accompanying open reading frames, was used to transform the cry$^-$ strain of Bt (*cryB*) as previously described [39]. Cry2Aa was purified from the crystalline inclusions produced in the cells. Inclusions were harvested by cell lysis and centrifugation. Crystalline inclusions were washed repeatedly with 0.5 N NaCl to remove proteases and were transferred to buffer (10 mM Tris-HCl, 1 mM EDTA [pH 8.0]) with 2% mercaptoethanol. Titrating the pH to 10.5, using NH_4OH, solubilized the protein from the crystalline inclusion bodies. The protein was purified by Sephacryl S300HR column chromatography as described [40] and concentrated by ultrafiltration to 10 mg ml^{-1}.

Crystallization and Structure Determination

For recrystallization, hanging drops of the resulting concentrated protein (10 μl concentrated protein buffered as described above) were equilibrated against wells that contained Tris buffer (10 mM Tris-HCl, 1 mM EDTA [pH 8.0]). Crystallization was induced by the gradual shift to neutral pH as the mobile NH_3 diffused from the drops. Crystals were transferred to storage buffer (50 mM PIPES, 250 mM NaCl [pH 6.8]) with 2% mercaptoethanol. The resulting crystals are in spacegroup $P4_32_12$; unit cell constants a = 85.6 Å, c = 163.9 Å. They have one monomer in the asymmetric unit, an

estimated 34% solvent content, and diffract to ~3.0 Å using Cu K_α X-rays from a rotating anode generator and to 2.0 Å at a synchrotron source after flash freezing.

For the collection of data at 100K, the crystals were transferred in three steps to a final 20% solution of cryo-protectant (a 1:1 mixture of 1,2-propane diol and glycerol) and storage buffer and flash frozen in a cold nitrogen stream. X-ray diffraction data were collected at SSRL beamline 7.1 using a wavelength of 1.08 Å. Intensity data were integrated, scaled, and merged using HKL [41]. The overall Wilson B factor (3.0 Å < d < 2.2 Å) was 14 $Å^2$.

De novo phasing was achieved using multiple isomorphous replacement after attempts to find a molecular replacement solution to the phase problem employing the available coordinates of Cry3Aa and Cry1Aa were unsuccessful. The heavy atom derivatives (Table 1) were solved from difference Patterson maps as displayed using XtalView [42]. Difference Fourier inspection for minor sites and refinement of the heavy atom positions, occupancies, and B factors was completed in PHASES [43]. The resulting protein electron density map was subjected to solvent-flipping density modification, as implemented in Solomon [44]. The helical bundle was apparent in 5 Å maps; at 3 Å resolution, the correct enantiomorph was clear from its stereochemistry. Using Cry1Aa as the initial building template, polyalanine versions of the helical and jellyroll domains were manually positioned using O [45], and the fit was optimized using the real-space refinement package ESSENS [46]. Positional and simulated annealing refinement were carried out using the maximum likelihood target of XPLOR 3.85× [47].

Acknowledgments

We thank D.H. Dean, E.A. Zhukovsky, J. Finer-Moore, and R.J. Fletterick for helpful discussions during the course of this investigation. We thank V. Ramalingam for assistance in crystallization and data collection and G.K. Powell for providing us with the Cry2Aa clone. This work is based upon research conducted at SSRL, which is funded by the Department of Energy, Office of Basic Energy Sciences. This work was supported by the National Institutes of Health (GM-244485 to R.M.S).

Received: December 27, 2000
Revised: April 4, 2001
Accepted: April 6, 2001

References

1. Schnepf, E., et al., and Dean, D.H. (1998). *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol. Mol. Biol. Rev. *62*, 775–806.
2. Tojo, A., and Aizawa, K. (1983). Dissolution and degradation of *Bacillus thuringiensis* δ-endotoxin by gut juice protease of the silkworm *Bombyx mori*. Appl. Environ. Microbiol *45*, 576–580.
3. Aronson, A.I., Han, E.S., McGaughey, W., and Johnson, D. (1991). The solubility of inclusion proteins from *Bacillus thuringiensis* is dependent upon protoxin composition and is a factor in toxicity to insects. Appl. Environ. Microbiol *57*, 981–986.
4. Choma, T., Surewicz, W.R., Carey, P.R., Pozsgay, M., Raynor, T., and Kaplan, H. (1990). Unusual proteolysis of the protoxin and toxin from *Bacillus thuringiensis*. Structural implications. Eur. J. Biochem. *189*, 523–527.
5. Hofmann, C., Vanderbruggen, H., Höfte, H., Van Rie, J., Jansens, S., and Van Mellaert, H. (1988). Specificity of *Bacillus thuringiensis* δ-endotoxins is correlated with the presence of high-affinity binding sites in the brush border membrane of target insect midguts. Proc. Nat. Acad. Sci. USA *85*, 7844–7848.
6. Wolfersburger, M.G., Hofmann, C., and Lüthy, P. (1986). Interaction of *Bacillus thuringiensis* with membrane vesicles isolated from lepidopteran larval midgut. Zbl. Bakt. Mikrobiol. Hyg. I. Suppl. *15*, 237–238.
7. Van Rie, J., Jansens, S., Höfte, H., Degheele, D., and Van Mellaert, H. (1989). Specificity of *Bacillus thuringiensis* δ-endotoxins: importance of specific receptors on the brush border membrane of the mid-gut of target insects. Eur. J. Biochem. *186*, 239–247.
8. Slatin, S.L., Abrams, C.K., and English, L. (1990). Delta-endotoxins form cation-selective channels in planar lipid bilayers. Biochem. Biophys. Res. Commun. *169*, 765–772.
9. English, L., Readdy, T.L., and Bastian, A.E. (1991). Delta-endotoxin-induced leakage of $^{86}Rb^+$-K^+ and H_2O from phospholipid vesicles is catalyzed by reconstituted midgut membrane. Insect Biochem. *21*, 177–184.
10. Schwartz, J.L., Garneau, L., Savaria, D., Masson, L., and Brousseau, R. (1993). Lepidopteran-specific crystal toxins from *Bacillus thuringiensis* form cation- and anion-selective channels in planar lipid bilayers. J. Membrane Biol. *132*, 53–62.
11. Harvey, W.R., and Wolfersberger, M.G. (1979). Mechanism of inhibition of active potassium transport in isolated midgut of *Manduca sexta* by *Bacillus thuringiensis* endotoxin. J. Exp. Biol. *83*, 293–304.
12. Harvey, W.R., Cioffi, M., and Wolfersberger, M.G. (1986). Transport physiology of lepidopteran midgut in relation to the action of *B.t.* delta-endotoxin. In Fundamental and Applied Aspects of Invertebrate Pathology, J.M. Vlak, D. Peters, and R.A. Samson, eds. (Wageningen, The Netherlands: Grafisch Berdrijf Ponsen and Looijen), pp. 11–14.
13. Knowles, B.H., and Ellar, D.J. (1987). Colloid-osmotic lysis is a general feature of the mechanism of *Bacillus thuringiensis* d-endotoxins with different insect specificity. Biochem. Biophys. Acta *924*, 509–518.
14. Wolfersberger, M.G. (1992). V-ATPase energized epithelia and biological insect control. J. Exp. Biol. *172*, 377–386.
15. Crickmore, N., et al., and Dean, D.H. (1998). Revision of the nomenclature of the *Bacillus thuringiensis* pesticidal crystal proteins. Microbiol. Mol. Biol. Rev. *62*, 807–813.
16. Losey, J.E., Rayor, L.S., and Carter, M.E. (1999). Transgenic pollen harms monarch larvae. Nature *399*, 214.
17. Van Rie, J., McGaughey, W.H., Johnson, D.E., Barnett, B.D., and Van Mellaert, H. (1990). Mechanism of insect resistance to the microbial insecticide *Bacillus thuringiensis*. Science *247*, 72–74.
18. McGaughey, W.H., Gould, F., and Gelertner, W. (1998). Bt resistance management. Nature Biotechnol. *16*, 144–146.
19. Yamamoto, T., and McLaughlin, R.E. (1981). Isolation of a protein from the parasporal crystal of *Bacillus thuringiensis* var. *kurstaki* toxic to the mosquito larva *Aedes taeniarhynehus*. Biochem. Biophys. Res. Commun. *103*, 414–421.
20. Donovan, W.P., Dankocsik, C.C., Gilbert, M.P., Gawron-Burke, M.C., Groat, R.G., and Carlton, B.C. (1988). Amino acid sequence and entomocidal activity of the P2 crystal protein. An insect toxin from *Bacillus thuringiensis* var. *kurstaki*. J. Biol. Chem. *263*, 561–567.
21. Yamamoto, T., and Powell, G.K. (1993). *Bacillus thuringiensis* crystal proteins: recent advances in understanding its insecticidal activity. In Advanced Engineered Pesticides. L. Kim, ed. (New York: Marcel Dekker), pp. 3–42.
22. Kota, M., Daniell, H., Varma, S., Garczynski, S.F., Gould, F., and Moar, W.J. (1999). Overexpression of the *Bacillus thuringiensis* (Bt) Cry2Aa2 protein in chloroplasts confers resistance to plants against susceptible and Bt-resistant insects. Proc. Nat. Acad. Sci. USA *96*, 1840–1845.
23. English, L., et al., and Slatin, S.L. (1994). Mode of action of CryIIA: a *Bacillus thuringiensis* delta-endotoxin. Insect Biochem. Mol. Biol. *24*, 1025–1035.
24. Widner, W.R., and Whiteley, H.R. (1990). Location of the dipteran specificity region in a lepidopteran-dipteran crystal protein from *Bacillus thuringiensis*. J. Bact. *172*, 2826–2832.
25. Liang, Y., and Dean, D.H. (1994). Location of a lepidopteran specificity region in insecticidal crystal protein CryIIA from *Bacillus thuringiensis*. Mol. Microbiol. *13*, 569–575.
26. Widner, W.R., and Whiteley, H.R. (1989). Two highly related insecticidal crystal proteins of *Bacillus thuringiensis* subsp. *kurstaki* possess different host range specificities. J. Bacteriol. *171*, 965–974.
27. Audtho, M., Valaitis, A.P., Alzate, O., and Dean, D.H. (1999). Production of chymotrypsin-resistant *Bacillus thuringiensis* Cry2Aa1 delta-endotoxin by protein engineering. Appl. Environ. Microbiol. *65*, 4601–4605.
28. Li, J., Carroll, J., and Ellar, D.J. (1991). Crystal structure of insec-

ticidal δ-endotoxin from *Bacillus thuringiensis* at 2.5 Å resolution. Nature *353*, 815–821.
29. Grochulski, P., et al, and Cygler, M. (1995). *Bacillus thuringiensis* CrylA(a) insecticidal toxin: crystal structure and channel formation. J. Mol. Biol. *254*, 447–464.
30. Lee, M.K., Young, B.A., and Dean, D.H. (1995). Domain III exchanges of *Bacillus thuringiensis* Cry1A toxins affect binding to different gypsy moth midgut receptors. Biochem. Biophys. Res. Commun. *216*, 306–312.
31. de Maagd, R.A., et al., and Bosch, D. (1996). Domain III substitution in *Bacillus thuringiensis* delta-endotoxin CrylA(b) results in superior toxicity for *Spodoptera exigua* and altered membrane protein recognition. Appl. Environ. Microbiol. *62*, 1537–1543.
32. Schwartz, J.L., Potvin, L., Chen, X.J., Brousseau, R., Laprade, R., and Dean, D.H. (1997). Single-site mutations in the conserved alternating-arginine region affect ionic channels formed by CrylAa, a *Bacillus thuringiensis* toxin. Appl. Environ. Microbiol. *63*, 3978–3984.
33. Von Tersch, M.A., Slatin, S.L., Kulesza, C.A., and English, L.H. (1994). Membrane-permeabilizing activities of *Bacillus thuringiensis* coleopteran-active toxin CrylIIB2 and CrylIIB2 domain I peptide. Appl. Environ. Microbiol. *60*, 3711–3717.
34. Foote, J., and Winter, G. (1991). Antibody framework residues affecting the conformation of the hypervariable loops. J. Mol. Biol. *224*, 487–499.
35. Wedemayer, G.J., Patten, P.A., Wang, L.H., Schultz, P.G., and Stevens, R.C. (1997). Structural insights into the evolution of an antibody combining site. Science *276*, 1665–1669.
36. Huang, F., Buschman, L.L., Higgins, R.A., and McGaughey, W.H. (1999). Inheritance of resistance to *Bacillus thuringiensis* toxin (Dipel ES) in the European corn borer. Science *284*, 965–967.
37. Hilbeck, A., Moar, W.J., Pusztai-Carey, M., Fillipini, A., and Bigler, F. (1999). Prey-mediated effects of Cry1Ab toxin and protoxin and Cry2A protoxin on the predator *Chrysoperla carnea*. Entomol. Exp. Appl. *91*, 305–316.
38. Bowen D., et al., and ffrench-Constant, R.H. (1998). Insecticidal toxins from the bacterium *Photorhabdus luminescens*. Science *280*, 2129–2132.
39. Sasaki, J., et al., and Yamamoto, T. (1996). Insecticidal activity of the protein encoded by the cryV gene of *Bacillus thuringiensis kurstaki* INA-02. Curr. Microbiol. *32*, 195–200.
40. Yamamoto, T. (1989). Identification of entomocidal toxins of *Bacillus thuringiensis* by high-performance liquid chromatography. ACS Symposium Series, *432*, 46–60.
41. Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.
42. McRee, D. (1993). Practical Protein Crystallography. (San Diego, CA: Academic Press).
43. Furey, W., and Swaminathan, S. (1997). PHASES-95: A program package for the processing and analysis of diffraction data from macromolecules. In Methods Enzymology C. Carter and R. Sweet, eds. (Orlando FL: Academic Press), pp. 307–326.
44. Abrahams, J.P., and Leslie, A.G. (1996). Methods used in the structure determination of bovine mitochondrial F-1 Atpase. Acta Crystallogr. D *52*, 30–42.
45. Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.
46. Kleywegt, G.J., and Jones, T.A. (1997). Template convolution to enhance or detect structural features in macromolecular electron-density maps. Acta Crystallogr. D *53*, 179–185.
47. Brünger, A.T. (1993). X-PLOR Version 3.1 a System for X-ray Crystallography and NMR. (New Haven, CT: Yale University Press).
48. Ferrin, T.E., Huang, C.C., Jarvis, L.E., and Langridge, R. (1988). The MIDAS display system. J. Mol. Graph. *6*, 13–27, 36–37.
49. Nicholls, A. (1992). GRASP Manual. (New York: Columbia University).
50. McDonald, I.K., and Thornton, J.M. (1994). Satisfying hydrogen bonding potential in proteins. J. Mol. Biol. *238*, 777–793.
51. Barton, G.J. (1993). ALSCRIPT: A tool to format multiple sequence alignments. Protein Eng. *6*, 37–40.

Accession Numbers

The coordinates and structure factors for Cry2Aa have been deposited with the Protein Data Bank (accession code 1I5P).